Chemical Structure Association Trust

ISSUE 5

NEWSLETTER

WINTER 2003

Personalia

The ACS Division of Chemical Information has announced that Professor Peter Johnson from Leeds University, UK, will receive the 2004 Herman Skolnik Award for his outstanding contribution to chemical information science.

The ACS Award for Computers in Chemical and Pharmaceutical Research, sponsored by Accelrys, will be awarded to Professor W. Graham Richards of the University of Oxford, UK.

David Spellmeyer is now at IBM, Almaden.

Steve Heller has a granddaughter, Samantha Rose, born weighing 8lbs 3oz on September 11, 2003. He now has 2 grandsons and a granddaughter.

Helen Cooke, formerly Schofield, will be leaving UMIST in January to work at GlaxoSmithKline in Philadelphia, as a training and development consultant in the Information Management, Education and Support Group.

The 2003 ExemplarChem prize sponsored by the CSA Trust was awarded to Helena Alsbury and Andrew Robertson, of the Department of Chemical Engineering, Loughborough University, Loughborough, UK. The title of their entry was 'Production and evaluation of photocatalytic titanium dioxide coatings on glass tubes.'

Sponsorship Secretary John Holliday will maintain the register of donors. Please send donations to him: Dr John Holliday, Department Of Information Studies, University of Sheffield, Sheffield, S10 2TN, UK. Tel: +44-(0)114-222-2685; e-mail: J.D.Holliday@sheffield.ac.uk

---

Chemical Structure Association Trust Annual General Meeting and Dinner

The CSA Trust AGM will be held on Monday 1 December 2003 at 4pm in the Roscoe Room at the Society of Chemical Industry (SCI), 15 Belgrave Square, London SW1X 8PS. The plan is to end the meeting at 6pm at the latest. After the AGM, the Annual Dinner will be held in the Council Room at the SCI at 6.30pm for 7pm. There will be time for attendees to use the changing facilities and rest rooms at the SCI before the meal. The total cost of the Dinner will be £47 per head. This includes a choice of pre-dinner drinks (sparkling white wine, red and white wine, orange juice, mineral water), four course menu (soup, turkey etc. with vegetables, Christmas Pudding with brandy sauce, coffee and mince pie), plus red and/or white wine with the meal.

To book for the dinner and for details of how to get to the SCI, please contact Dr J. Clive Weeks, CSA Trust Secretary, tel: +44-(0)1252-674890; e-mail: drcliveweeks@aol.com

Visit the CSAT website at http://www.csa-trust.org

Inside...

Personalia

CSA Trust AGM and Dinner

Nîmes Conference Report

ACS Meeting Report

Leslie Bretherick

Dalton Celebrations

Jane Whittall

Events 2003–2004

Product News

Ernie Hyde Award through the Ages

Contact Details

Next Copy Date

International Chemical Information Conference, Nîmes

The 15th International Chemical Information Conference was once again held in Nîmes, which is not only notable for its splendid Roman remains, but is also, so we were told, the fifth safest city in France. The conference opened on Sunday 19th October with the traditional cocktail and buffet sponsored by CAS.

Monday started with a keynote talk from David Evans, Clairvoyance Corporation, on 'Clustering and Information Management'. We are entering a new era where the limitations are human factors, not the technology. Pseudo Relevance Feedback (PRF) can improve retrieval: taking the top documents from a hit list, extracting terms from those documents, and using the extracted terms as additional search terms. Clustering gives even better results. Automatic clustering is significantly improved by a few minutes of human intervention per search.

The first session was on Changing Technology in Information Services. Thomas Lorenz discussed the impact of changing technology at BASF. There are two main systems in operation: BLITZ (the in-house registry system) and SIRIUS, a text-oriented system. Users require a single interface to access all internal and external business and scientific information. To provide such integration requires both technology and content, with a lot of interdisciplinary cooperation, but at BASF they conclude that the effort required to create a fully integrated system is not justified.

Engelbert Zass, from ETH Zurich, looked back at changing technologies. In 1989, end users never did any searching, but in 1998, the library offered eight databases at the bench, and it became an 'Information Centre' offering assistance, training and education for users. Now, the library runs seven servers with four operating systems (formerly operated by the Computer Centre), and librarians, chemists and biologists are all available to help users. Users are beginning to influence future database developments, e.g. in Beilstein.

The experiences of CAS were outlined by Robert Swann. Technology drives change and change drives technology, but it is the interaction of technology and content that leads to innovation. There is a constant demand for cheaper, simpler, faster systems. Robert referred to two IBM Projects: the Millipede project and the Blue Gene project. Now 93% of information is 'born digital' which affects the way companies interact with users.

Five product reviews followed, out of a total of 18. A key feature of the Nîmes conference is the exhibition,

The CSAT Trustees having a serious meeting!

sited close to the lecture theatre, and exhibitors are each given an 8-minute (strictly timed!) slot to introduce their product.

Session two of the conference covered New Computing and Analytics, and Karl Harrison from Oxford University outlined the potential of Grid Computing on a massive scale. Grid computing can involve the linking of supercomputers, as in the UK e-Science programme, or involve a massive grid of desktop computers, such as the SETI@ home project run by Berkeley University in the US which searches the skies for extraterrestrial intelligence. The National Foundation for Cancer Research Centre for Computational Drug Design in Oxford has started a project to use a grid of more than 300,000 desktop computers, in which the screensavers are used to determine the binding possibilities of protein and ligand. This uses a total of 60–100 teraflops of processing power, whereas the largest single machine in the world is only(!) 35 teraflops. The technology is also being used to look for potential drugs to combat anthrax and smallpox terrorist threats. The US government are testing the hits found in the anthrax project.

The next talk was on Patent Data and Competitor Intelligence. The original speaker was unable to attend and was replaced at very short notice by Charles Huot from TEMIS. Not surprisingly, Charles' talk came over as an extended Product Review of the TEMIS software editor, which uses text mining to convert text data into business value using automatic clustering and classification. Bob Stembridge from Thomson Derwent continued the patents theme with his talk 'Patent Analysis, Tips, Traps and Techniques'. Patent analysis is cheap, easy and available, but it does require a good search strategy, good cleaned data and an understanding of data sources, codes and classifications. Using patent analysis can help with assessing the quality and quantity of competitors' work and to show where a company is specialising.

Andrej Budo-Marek from Xerox Global Services finished the day with 'Trends in intellectual asset management'. The value of paper-based information is underestimated, but there is the problem of making it digital. Performing a search and distributing information is time-consuming, but the most valuable part of the process is analysing the data and delivering recommendations. Intellectual asset management always involves human intervention.

On Monday evening, delegates enjoyed the Conference Cocktail sponsored by MDL Information systems, Elsevier and ChemWeb.

The question of whether or not we should pay for information was considered on the second day. Liz Blankson-Hemans, Quantum Dialog, gave a very clear presentation on 'Defining the value of Information: beyond ROI (Return On Investment)'. It is essential to be able to show the benefits of an information resource centre (IRC). It is not sufficient to say that the IRC is productive, saves time and is cost-effective, but specific examples must be given of how the IRC has solved business problems, e.g. to demonstrate that it has helped to bring a drug to market early.

William Woishnis from Knovel came back for his second presentation of the conference (he gave a Product Review on the previous day). This again was a late-replacement lecture, as the planned speaker had just a baby – which left us wondering why this could not have been predicted somewhat earlier! Knovel was described as an aggregator portal dealing particularly with handbooks and reference materials, where the aim is to make that information more valuable than the original source. As was stressed many times in the conference, it's quality, not quantity that matters. At Knovel much of the work is done manually.

Minoo Philipp, from Henkel, then described their fully self-supporting information centre, where they buy the information and sell it on to their clients within the organisation. They can sell only on the basis of the quality offered. Users want to search with one common language over several databases, possibly back over a long period of time. The essential ingredients for making ends meet are good indexing (or clustering), timeliness, user training, customer support, good ways of presenting the search results and careful pricing.

Another view from the value-added provider came from Ricardo Vieira of Micropatent. Ricardo was yet another replacement speaker, but he came up with a memorable talk about water! He compared paying for information to buying bottled water – it's expensive to 'bottle' information and the vendors need to charge. If you use a free service, you have no guidance as to what you are searching and you cannot make any demands of the service. Users who pay have 'rights' and can make demands on vendors. Competition is important and it will lead to competitors striving to improve services. At the end of the talk, EPO (whose services are available free) commented that 10–15 years ago their discussions were aimed at those who pay, but then it was the users who requested free information.

The next session covered 'Search and Retrieval'. The first paper was from Chahab Nastar, LTU Technologies, France, where they are

working on content-based image retrieval. IBM pioneered the technology about 10 years ago with QBIC (Query By Image Content) but there are few companies working in this area. Although images are spontaneously recognised by humans, it is much more difficult for computers, with the lack of alphabet, taxonomy or mathematical properties in images. Image content can be described using keywords added manually (but this is subjective) or by similarity (detected automatically). Similarity retrieval can be used to generate keywords, using feature extraction and encoding. Content-based image retrieval is important in searching patent and trademark databases for brand protection and copyright infringement.

The morning session ended with Stephen Arnold, AIT, talking about Social Software and a new search paradigm. Social Network Technology is changing the way we communicate and the number of people communicating. One billion SMS messages a day are transmitted in Europe. Blogs are being used to create Newstreams (see Henry Rzepa's article in the last Trust Newsletter). We must look beyond the enabling technologies used on the web (spiders, indexing and search) to the use of analytical technologies such as filtering, counting, clustering and data mining. Social software, which includes messaging technology (for example, e-mails, SMS, video), context technology (e-commerce, live chat, smart agents) and community technology (groups, blogs), will dramatically change the way in which we communicate. Content needs to be redefined by all the new means of communication.

'Publishing at a cross-roads: who is going to be doing what?' was the title of the contribution from Stephen Bachrach of Trinity University, Texas. Chemistry publication is still driven by the journal. The use of the internet for publishing is a worry for publishers (that they will no longer be needed) and a worry for scientists (that it will be the end of peer review). Publishing on a personal website provides no archiving, no peer review and no adequate search procedures. Alternatively, scientists can publish on a preprint server, but the ChemWeb preprint server has had little impact. In web publishing, there should be a trend towards publishing fewer, better articles, with as much data as possible, such as molecules and spectra in Chime, spreadsheets in Excel, movies in Real Player and large data sets. XML is the enabling technology, and standards such as CML and IChI are needed. This can only be achieved by close cooperation between publishers and scientists.

Michel Vajou from M.V. Etudes et Conseil, France gave a presentation on 'Concentration and integration in the information industry: towards vertical monopolies in the provision of online professional information'. For more than 10 years, there have been many mergers and acquisitions, leading to very large leaders in the field, such as Elsevier. Integration has been difficult, as companies have had difficulty realising the economies of scale, and the benefits of being global. The provision of 'vertical' products and services suffered from incompatibility. A long-term profitable relationship with the desktop users is essential for survival.

The final talk on Tuesday came from Peter Shepherd of COUNTER, UK, on 'Standards for the recording and exchange of online usage data: towards fairness for publishers and customers'. Usage statistics, needed for making improvements, experimenting, planning, new pricing models, analysis of demographics, and so on, must be credible, compatible and consistent. COUNTER is a not-for-profit company that has been set up to provide a single Code of Practice. The first version was published in January 2003. The website (http://www.project counter.org) has a list of the companies that are COUNTER compliant.

Tuesday evening provided a new, and very welcome, change from the previous Nîmes conferences, in that there was a conference dinner and entertainment, thanks to the generosity of MicroPatent. We were transported to the Carmague (on buses sponsored by Questel.Orbit), to see a demonstration of training young bullocks for bull-fighting. It was a chilly night and we were pleased to arrive at a really splendid barn at Le Mas de Peint for an excellent meal, and further entertainment by a group of instrumentalists and two flamenco dancers.

Session 6 on 'The New Patent World' included a talk by Gerard Giroud of the European Patent Office on 'Collection of chemical structure data during patent filing'. EPO are developing an electronic filing tool for chemical structures. The system will accept structures drawn in standard structure editing programs, and then the structure will then be verified and prepared for filing. The prototype is due in December. It is hoped that this will significantly increase the number of users who file patents electronically.

David Dickens from Questel.Orbit, US, Pierre Buffet from Questel. Orbit, France and Yuji Takashima from PATOLIS, Japan then presented a guide to Japanese patent information. Japan is the world's number one patent producer, accounting for 40% of the world's patents, but the majority are not filed outside Japan. There is increasing coverage of Japanese patent information in English, via various sources, but there are limitations in all of the services. PATOLIS-e is by far the most comprehensive.

Tuesday night's flamenco dancing included a guest appearance by Bill Town

Stephen Stein opened the final session on 'Classification and Categorisation', with his lecture on the IUPAC Chemical Identifier, IChI. IUPAC nomenclature is unsuitable for the digital processing of chemical structures and the IChI project, started three years ago, aimed to develop a set of algorithms to uniquely identify a compound, i.e. a digital signature. Problems included compounds with ill-defined connectivity, and chemists using differing conventions. These were overcome by normalising and simplifying the structure, then adding layers to describe such things as tautomers, stereochemistry and isotopes. The identifier, represented in XML, consists of the basic structure and the additional layers. A working version is expected by the end of 2003. Now it is important that the database producers and software companies adopt the identifier.

Caspar Fall and Patrick Fievet from ELCA Informatique and WIPO, Switzerland, gave a joint presentation about the WIPO system to assist users to categorise patent documents in the International Patent Classification (part of the CLAIMS project). Patent categorisation requires machine learning techniques to recognise the topics. The statistical distribution of the words in a document is established, and then a set of training documents that already have good IPC codes is used. The prototype, based on the Winnow algorithm, is giving excellent results at class level in English, French, Russian and German, when compared with manual assignment. Further work is needed at sub-class and main group level. The algorithm is designed to assist manual categorisation, not to replace it.

Anthony Rowe from Imperial College talked about the Wellcome Trust project on domain-mapping systems to integrate biological and chemical analysis. Research projects tend to be based on a specific disease or a protein family, and multidisciplinary information is required. The Wellcome Trust project aims to combine reference information with the underlying experimental data dynamically, using text mining and natural language processing for the literature database, and classification and clustering for the chemical and biological databases.

The final lecture, by Kirill Degtyarenko from the European Bioinformatics Institute in Cambridge, discussed the vocabularies and ontologies for bioinformatics. Ontology is a controlled vocabulary of terms which have defined logical relationships to each other. Several taxonomies have been developed in biology, (e.g. for the NCBI database, and the Gene Ontology (GO) consortium), and there are various resources being developed at the European Bioinformatics Institute, but there is a lack of an authoritative database of biochemical compounds in the public domain. The EBI is developing a dictionary of Chemical Compounds of Biological Interest (ChEBI). This will provide a standard for biochemical terminology and for structures, using the IChI, described earlier.

The conference ended with the usual votes of thanks, but I would like to propose a vote of thanks to Harry Collier of Infonortics, who masterminds this excellent conference and ensures that everything proceeds according to plan. There was a good mix of papers – with over 200 people attending the conference, from 16 different countries and many different backgrounds – it is essential to cover a wide area so that there is something for everyone. Copies of the lectures are available at http://www.infonortics.com/chemical/ch03/03chempro.html.

Next year, the conference returns to Annecy, from 17th–20th October.

Janet Ash

226th American Chemical Society National Meeting – Fall 2003 New York, September 7–11

Wendy Warr and Bill Town

Attracting an audience of about 14,000 people, to hear nearly 7000 papers and visit 528 exhibition booths, this was a meeting of some size and complexity, typical of most six-monthly National Meetings of the American Chemical Society.

The last New York meeting was in Autumn 1991, in an assortment of hotels. Much time was wasted pounding the pavements (or sidewalks, if you will) between sessions in different locations. This time many of the technical sessions and meetings, and the Exhibition, took place in the Jacob K. Javits Convention Center and regular shuttle buses connected the Center to a selected number of hotels. Nevertheless, the New York traffic still prevented the dedicated attendee from cutting and pasting from programs in different locations.

The dilemma was made even worse on some days, when not only did an interesting COMP session clash with a CINF one, but parallel programming within CINF led to a tantalising toss of a coin. (To the cognoscenti, CINF is the abbreviation for the Division of Chemical Information and COMP is the abbreviation for the Division of Computers in Chemistry). 'Technical intelligence: primary source analysis' for example, clashed with 'Challenges for the chemical sciences in the 21st century: information and communication'. Worse still, 'Advances in reaction searching' clashed with 'Emerging trends in discovery data integration'.

Frank Allen (left) receives his Herman Skolnik Award from Chuck Huber

Between us, we managed to cover some of the papers in those sessions and many of those in symposia entitled: 'Crystallographic databases and their applications'; 'Novel database and knowledge mining techniques'; 'Standards for chemistry informatics'; 'Technical intelligence. Lessons from competitive intelligence'; and 'Building the virtual chemistry library: e-books and e-journals'. Detailed summaries will appear in *Chemical Information and Computation* in January 2004.

The Herman Skolnik Award of the ACS Division of Chemical Information, recognising outstanding contributions to and achievements in the theory and practice of chemical information science, was presented to Frank Allen. Frank has driven the development of much of the software for data acquisition and validation, and for the retrieval and analysis of information in the Cambridge Structural Database (CSD), and has also pioneered many of its research applications.

The award symposium, on crystallographic databases and their applications, began with two speakers from CCDC. Frank Allen described the CSD and its research applications in structural chemistry, while Robin Taylor discussed data mining of crystallographic databases as an aid to drug design. The morning session ended with two speakers from the Protein Data Bank (PDB): Helen Berman described the evolution of the PDB and Phil Bourne presented the PDB as a research tool.

There was a change of focus in the afternoon session, with Carol Brock discussing when fractional crystallisation can be expected to fail, using information from the CSD, Guy Orpen presenting applications of the CSD to molecular inorganic chemistry, and John Rodgers (a former CCDC staff member) presenting his work on materials informatics and knowledge acquisition for materials design. The session ended with G. Ceder, with a paper on first principles calculated databases for the prediction of intermetallic structure.

The day-long symposium database and knowledge mining techniques was particularly interesting. Norah MacCuish made perceptive observations on clustering ambiguity, and Susan Bassett described Bioreason's approach to automating rule discovery from data. Robert Sheridan of Merck has mined MDDR using TIMI, a method of relating words and chemical structures in a set of documents. Xiao Lewell of GlaxoSmithKline described a web-based tool to help in ring replacement strategies. Liz Colbourn of Intelligensys (an ex-ICI colleague) showed remarkable aplomb when smoke started to rise from the overhead projector during her presentation on knowledge mining in formulation databases, and continued without visual aids, but with clarity and good humour.

Two papers were withdrawn from the symposium on standards for chemistry informatics, leaving the session dominated by speakers from NIST. Steve Stein talked about the IUPAC chemical identifier and an IUPAC project intended to create a set of fully parsed, XML data dictionaries based on IUPAC standard glossaries (the 'color books'). MDL was well represented in the session on reactions, with three papers, including one from the Trust's Chairman, Guenter Grethe. Lingran Chen gave a significant paper on MDL's reaction indexing and searching methodologies, past and present, incorporating some details, as yet unpublished, of the current algorithms. Harold Varmus spoke in a Chemical Education symposium entitled 'Communicating chemistry revisited'. A report of his paper on open access publishing, written by Bonnie Lawlor, is available to CINF members.

The meeting proved a good occasion for meeting colleagues both old and new. Stu Kaback, a former Herman Skolnik awardee and staunch supporter of CINF Division for many years, retired recently, but returned for a CINF social event in New York. It was a pleasure to see Sandor Barcza again asking perceptive questions in technical sessions. Richard Williams, too, returned to the scene, with a new affiliation: Williams Information Services. Seven 'old' faces (but looking never a year older) gathered for a CSA Trustees meeting one lunchtime: Guenter Grethe, Andreas Barth, Bonnie Lawlor, Peter Nichols, Peter Rusch, Bill Town, and Wendy Warr. Incidentally, the CSA Trust will be co-sponsoring a symposium on grid computing at the ACS Spring 2004 meeting in Anaheim: mark your diaries.

Beyond the parochial events of our favourite divisions, the American Chemical Society as a whole celebrated the eightieth anniversary of *Chemical and Engineering News* and the publication of the 125th volume of the *Journal of the American Chemical Society*, with a special day-long symposium, and a black-tie gala reception and dinner attended by 700 people and sponsored by 39 organisations.

A memorable meeting ended on a sombre note: the last day was the second anniversary of the 9-11 tragedy. Co-resident with us at the Marriott was a large contingent of British bobbies. Seeing them in dress uniform, with immaculate white gloves, on September 11, aroused a mixture of emotions, and no-one could have failed to be moved by the voice of the soloist singing 'God Bless America' shortly before 9am: the sound echoed through the hotel atrium, up and beyond the thirtieth floor. Better to regard it as the voice of hope, beckoning us to the next ACS meeting in the city in 2010.

Leslie Bretherick

Leslie Bretherick, whose *Handbook of Reactive Chemical Hazards* made him a well-known name in the field of chemical safety, died this year. Butterworth Heinemann published the first paper edition in 1975 and in 1999 the initial online version was launched on Chemweb. The idea of producing the Handbook came about when Bretherick was working in the chemical industry and a colleague was killed in an explosion. Leslie vowed to devote the rest of his career to collating information that might save lives in the future, and continued his work for many years, even when his eyesight was failing. He was a consultant safety editor to other publications, such as the *Dictionary of Organic Compounds*. In 1988, Leslie received the ACS Chemical Health and Safety Award for his contributions to chemical health and safety.

Celebrations of the Bicentenary of John Dalton's Atomic Theory

An exhibition to celebrate the Bicentenary of Dalton's Atomic Theory is being held in the Manchester Museum at the University of Manchester from October 14 2003 to January 5 2004. The exhibition has been prepared by the John Rylands University Library of Manchester in conjunction with the Manchester Literary and Philosophical Society.

The opening ceremony, sponsored by ACS, took place on October 13. Speeches were given by Mr Bill Simpson, Librarian JRULM; Dr Diana Leitch, Exhibition Organiser and Deputy Librarian; Professor Sir Harry Kroto, FRS, President RSC and Professor Dave Garner, FRS, who heads up the Dalton Division of RSC. Kroto explained that Dalton was the first nanoscientist, and in the afternoon Kroto and Garner had been talking about buckyballs to 300 schoolchildren, who are Manchester United fans.

Elegant library staff in period costume at the opening ceremony

John Blundell-Ellis (who runs PsiGate from the JRULM) was dressed as Dalton (and he had to stand all night, probably because of the tight breeches that did not encourage more sedentary poses!) Library ladies were dressed in elegant Regency costumes. The toast to John Dalton was given by the President of the Manchester Literary and Philosophical Society (of which Dalton was the first president). The Adlington Folk Dancers contributed to the evening with dances of the period. The recently refurbished Manchester Museum was thanked for providing space for the Exhibition.

Many of Dalton's publications are on display at the exhibition, including his first published work, *Meteorological Observations and Essays*, 1793. He kept a meteorological diary for most of his life, and his aim in publishing these notes was to help people to use thermometers and barometers, and to understand the results. He had also discovered 'the relation of the aurora borealis to magnetism' and wanted to present these conclusions to the public. The pages shown are a record of the temperatures at Kendal and Keswick in Cumbria, UK. Dalton comments that the thermometer at Kendal was 'under the shade of a pretty large gooseberry tree, facing north'; that 'at Keswick was in an open situation, also facing north, so that it was likewise not affected by direct sunshine'.

John Dalton was born in Eaglesfield, Cumberland (now Cumbria) in 1766. His father was an impoverished Quaker weaver, and Dalton's early education was in the village school. By the time he was 12, Dalton himself was teaching other children in the village, and for a short time he worked in the fields for local farmers.

In 1781 Dalton went to Kendal as teacher in a Quaker school there, and when the schoolmaster retired in 1785 Dalton and his brother Jonathan took over the school. Quakers from the Lake District were represented at the opening ceremony.

While in Kendal, his interest in meteorology developed and he began to keep a meteorological journal, which he maintained all his life. He was befriended in Kendal by the blind naturalist John Gough; he began to collect plants and formed a herbarium. In 1787–1788 he gave a series of public lectures in Kendal on natural history.

In 1793 Dalton moved to Manchester. His original reason for coming to the city was to take up the post of tutor in mathematics and natural philosophy in New College, Manchester, an academy set up to provide higher education for Dissenters who could not attend the existing English universities. In October 1794 Dalton was admitted as a member of the Manchester Literary and Philosophical Society, and later that month read his first paper to the Society on 'Extraordinary Facts Relating to the Vision of Colours'.

Dalton had discovered that his perception of colours differed from that of most people when he was looking at a flower which most people saw as pink, but he regarded as blue. Colour-blindness is sometimes called 'daltonism' after Dalton's research into this area.

In 1799 he read a second paper to the Literary and Philosophical Society, on Rain, Dew and Evaporation. Also in 1799 New College removed to premises in York, and Dalton henceforward supported himself by freelance teaching of

scientific subjects. In 1800 he was elected as secretary of the Literary and Philosophical Society.

In 1801 Dalton read an important paper to the Society 'On the Constitution of Mixed Gases' which placed meteorology on a scientific basis, and also in this year his book on English grammar was published. In 1803 his paper read to the Society 'On the Absorption of Gases by Water' first presented his atomic theory to the public. In 1804 Dalton developed his ideas in a series of lectures to the Royal Institution in London, and in 1807 gave a series of lectures in Edinburgh. In 1808 Dalton brought his ideas together in 'A New System of Chemical Philosophy', and also in 1808, he became vice-president of the Literary and Philosophical Society.

Dalton gained many distinctions. In 1817 he became President of the Literary and Philosophical Society, a post he held until his death. In 1825 he was awarded the first annual prize of the Royal Society, and in 1831 was present at the first meeting of the British Association for the Advancement of Science at York. He was presented with an honorary D.C.L. at the meeting of the British Association in Oxford in 1832, and Edinburgh granted him an honorary LL.D. in 1834. In 1842 the British Association met in Manchester, but by this time Dalton was too frail to undertake the presidency, and was appointed vice-president. Dalton died in Manchester in 1844.

There were two weeks of Dalton celebrations in Manchester, and the exhibition, which is well worth a visit, is open until January 5 2004. For further details, see http://rylibweb.man.ac.uk/dalton/index.html

Thanks to JRULM for permission to publish, and to Wendy Warr who described the opening ceremony

Jane Whittall (née Gaworska)

Jane Whittall who died in June 2003 at the age of 53 was a career information professional to her fingertips. Her working life was spent in the pharmaceutical industry. Following a degree in Biochemistry and an MSc in Information Science, Jane's first job in information science was within Beecham Pharmaceuticals' corporate division. She was quickly enticed to join their R&D Information Service. Beecham merged with SmithKline Beckman to become SmithKline Beecham which much later merged with GlaxoWellcome to form GSK. On each occasion Jane's responsibilities grew and, through 2001, as Director of Published Information for GSK's R&D Group, she built the team which defined GSK's scientific literature and news product portfolio and the mechanisms for its delivery across GSK's global organisation. Last year Jane took on her most challenging assignment: the 'Library of the Future' project. The legacy of this work will stand as a tribute to Jane for some considerable time.

Through her membership of IMPI (Information Managers in the Pharmaceutical Industry) and of the PDR (Pharma Documentation Ring) she was a powerful and engaging influence on professional colleagues. Jane joined the PDR as the SKB member in 1992 as the multinational membership expanded. Her experience and influence in the publishing industry stimulated her strong involvement in the successful PDR/STM model e-journal license project. The standard draft contract for the acquisition of electronic journals is used by many pharma companies as a starting point for negotiations with the publishing industry. It was a significant achievement and has underpinned the negotiations that GSK now has with publishers.

Jane also made a major contribution to the Chemical Structure Association in the 1980s when she was engaged in the WLN movement and the move towards improved chemical structure computerisation. A particular highlight was her involvement in the 1982 Exeter Conference, a full week devoted to all the significant developments of the time and attracting everyone who was everyone in chemical information management – Eugene Garfield giving the after dinner speech and the organising committee stimulated by Frances Barker to perform a sketch that borrowed from *The Hitch Hiker's Guide to the Galaxy* as their contribution to after dinner networking!

When the 'Library of the Future' project was undertaken, Jane's health had been under threat for some time. She had long suffered with multiple sclerosis, yet few of her colleagues knew. She had bleak times but you were unlikely to hear of them from her. After the diagnosis of pancreatic cancer, she was determined to cram as much as possible into the remaining two years of her life, with new projects at GSK, and the 44th PDR AGM at Newmarket which her superb organisation made a resounding success.

Away from work Jane loved camping, blackberrying, walking, architecture and things archaeological. She married Neil in 1982 and we extend sincere sympathy to Neil and their twin daughters, Jessica and Eleanor.

Jane was a remarkable person. Her bravery and fortitude were so impressive. She was bright, honest, loyal, pragmatic, funny, forward-looking, highly respected and widely loved. She was a brave, brave lady and will be sadly, sadly missed.

Sandra Ward

EVENTS 2003–2004

December

1 December: CSAT AGM and Dinner: see Page 1

2–4 December: Online Information 2003, Olympia, London, UK. Contact Fiona Ashton, Online Information, c/o Circulation Data Services (CDS), PO Box 6009, Thatcham, Berkshire, RG19 4TT, UK. Tel: +44-(0)1635-588863; fax: +44-(0)1635-868594; http://www.online-information.co.uk

2004

February

11–13 February: ADMET 1, Town & Country Hotel, Hotel Circle, San Diego, CA, USA. Contact Darrin Scherago, Scherago International, 11 Penn Plaza, Suite 1003, New York 10001, USA. Tel: +1-(212)-643-1750, ext. 20; fax: +1-(212)-653-1758; e-mail: darrins@scherago.com; http://www. scherago.com/admet/

March

7–10 March: International Patent Information Conference and Exposition, IPI-ConfEx, Lisbon, Portugal. Website: http://www.ipi-confex.com/

8–10 March: AccelrysWorld 2004, Hotel Del Coronado, 1500 Orange Avenue, San Diego, USA. Contact Beverly Batchelder, Accelrys, 9685 Scranton Road, San Diego, CA 92121, USA. Tel: +1-858-799-5000; fax: +1-858-799-5777; e-mail: accelrysworld@accelrys.com; http://www.accelrys.com/ accelrysworld/

March 28–April 1: 227th ACS National Meeting, Anaheim, CA, USA. Includes CSA Trust symposium on grid computing. ACS Meetings, 1155 16th St., NW, Washington, DC 20036-4899, USA. Tel: 1-800-227- 5558; fax: +1-(202)-872-6128; e-mail: natlmtgs@acs.org

April

18–21 April: Drug Discovery Frontiers in Cancer and Neuroscience. The Application of Chemistry to Diseases, Leavey Conference Center, Georgetown University Medical Center, Washington, DC, USA. Contact: Jessica Bauman, Drug Discovery Program, Research Building, Rm EP09, 3970 Reservoir Rd, NW, Washington, DC 20057, USA. Tel: +1-(202)-687-0151; fax: +1-(202)-687-0738; e-mail: drugdiscovery@georgetown.edu; http://drugdesign.georgetown.edu/ddc2004/index. htm

21–23 April: 3rd Joint Sheffield Conference on Chemoinformatics. CSA Trust/MGMS, University of Sheffield, Sheffield, UK. See Page 12

June

14–16 June: Chiral Europe 2004, Hilton Hotel, Mainz, Germany. Contact Dr Claire Davey, Scientific Update, Maycroft Place, Stone Cross, Mayfield, East Sussex, TN20 6EW, UK. Tel: +44-(0)1435-873062; fax: +44-(0)1435-872734; e-mail: claire@scientificupdate.co.uk; http://www.scientificupdate.co.uk/

July

11–14 July: 16th International Symposium on Chirality (ISCD) Kimmel Center for University Life at New York University, New York, USA. Contact Janet Cunningham, Barr Enterprises, PO Box 279, Walkersville, MD, 21793, USA. Tel: (+1)-301-668-6001; fax: (+1)-301-668-4312; e-mail: janetbarr@aol.com; http://www.nyu.edu/chirality/

18–21 July: 4th International Conference of the Chemical Societies of the South-East European Countries – Chemical Sciences in Changing Times: Visions, Challenges and Solutions, TMF, Karnegijeva 4, Belgrade, Serbia, 11000, Yugoslavia. Contact Ivanka Popovic, Serbian Chemical Society, Karnegijeva 4, Belgrade, Serbia, 11000, Yugoslavia. Tel: +381-11-33709-467; fax: +381-11-33709-467; e-mail: icosecs@elab.tmf.bg.ac.yu;http://www.shd.org.yu/icosecs4/

August

15–19 August: XVIIIth International Symposium on Medicinal Chemistry (ISMC), Swedish Academy of Pharmaceutical Sciences, Wallingatan 26 A, PO Box 1136, Stockholm, SE-111 81, Sweden. Fax: (+46)-820-55-11; info@ismc2004.dk/; http://www. ismc2004.dk/

22–26 August: ACS National Meeting, Philadelphia, USA. Contact details as for Spring meeting

September

5–10 September: Euro QSAR 2004, 15th European Symposium on Quantitative Structure Activity Relationships, Harbiye Museum and Congress Center, Istanbul, Turkey. Contact Turan Günes Blv. 28/3 B Blok 06550, Cankaya-Ankara-Turkey. Tel: +90-312-441-5686; fax: +90-312-441-5838; e-mail: armoria@euro-qsar2004.org; http://www.euro-qsar2004.org

October

4–5 October: Chiral USA 2004, Boston Marriott, Long Wharf, Boston, MA, USA. Contact Dr Claire Davey, details under June, Chiral Europe meeting

17–20 October: International Chemical Information Conference, Annecy, France

PRODUCT NEWS

FIZ CHEMIE helps find the web's hidden secrets

Hidden in the web are pages that normal search engines do not access, for whatever reason. It is, however, often those web pages that contain interesting information from universities and research institutes. Germany's 'Chemistry Information Centre' (FIZ CHEMIE Berlin) now offers three subject-specific internet search engines 'ChemGuide' (chemistry), 'MedPharmGuide' (medicine and pharmacology) and 'PublishersGuide' (scientific publishers). These powerful search engines access more than 7000 servers and will search these thoroughly, down to the last published page. All entry pages found by the Guides are evaluated by FIZ CHEMIE Berlin for relevance. With a single search query in one of these engines, one can thus search thousands of servers with approximately 15 million web pages.

Use of the search engines is free. An extensive retrieval language is available for search queries. As well as the usual Boolean operators, wildcards may be used and proximity searches are possible. Searches may also be undertaken automatically, this being especially useful for the permanent monitoring of developments in one's own research area or for complex search queries. In this case, the user sets up an 'SDI' (Selective Dissemination of Information), a search strategy which is carried out every time the database is updated and which informs the user only of relevant new hits. To carry out an SDI, it is necessary to carry out a retrospective search. These retrospective searches and the SDIs themselves are offered as annual subscriptions. The search engines are available at http://www.chemistry.de by clicking 'Databases'.

Science IP

A new CAS service, Science IP (www.scienceip.org) is intended to help information professionals, lawyers and others in the patenting and research and development communities without the time, staff or expertise to meet their scientific information retrieval needs. Services range from a relatively simple literature search for prior art to more involved projects that support the writing of patentability, freedom-to-practise/operate and validity opinions. In addition to industry experts in patents and scientific searching and retrieval, Science IP will also employ CAS scientists specialising in areas such as Markush chemical structures, nucleic and amino acid sequences, polymers, molecular biology, genetics, material sciences, medical devices, and pharmaceuticals. These Science IP advisors are experts in their respective fields who, as part of their CAS responsibilities, are constantly immersed in the patents and literature of their particular subject area. Their collective experience is available to assist the research staff in preparing the most comprehensive search strategies possible.

MDL Core Interface 1.1

MDL Information Systems, Inc. (MDL) has released MDL Core Interface 1.1, an upgraded version of its next-generation middleware technology for integrating scientific information and business processes. The new version includes substantial performance improvements while adding support for Sun Solaris 8 and 9 platforms and IBM WebSphere 5.0.1 application servers. Core Interface is the middleware server of the new MDL Isentris architecture platform. With its multi-tier framework and a set of core integration services, Core Interface enables discovery organisations to develop and deploy custom applications rapidly, and create integrated informatics environments to support research. To protect companies' existing data and ensure an uninterrupted work environment, MDL offers seamless migration to Isentris from MDL ISIS. MDL's website is at http://www.mdli.com/.

SYBYL 6.9.1/UNITY 4.4.1

New features with this software update from Tripos are FUGUE and FlexX-Pharm. FUGUE is designed to assist in recognising distant homologues (divergent evolutionary relationships) by sequence-structure comparison. This program provides structural and functional information about a target sequence and a basis for comparative protein modelling. FlexX-Pharm adds functionality to FlexX, enabling ligand-docking calculations to be guided through interaction (i.e., pharmacophore features) and spatial constraints. The update also includes code release on multiple platforms. Also available on Linux, SYBYL 6.9.1/UNITY 4.4.1 is the only complete computational chemistry suite available that can be used with, but does not rely on, mixed-platform or client-server architecture. Tripos has a website at http://www.tripos.com/

German patents coverage

Thomson Delphion now offers searching of full-text patent specifications from the German national collection. With main claims back to 1968, this is the most comprehensive set of German patent data available. The collection includes applications, granted patents, and utility models. More information can be found at http://www.delphion.com/press_releases/german.

Contact Details

Chair of Trustees

Guenter Grethe
Tel: +1-(510)-865-5152
e-mail: g.grethe@comcast.net

Treasurer

Geoff Downs
Tel: +44-(0)1909-515444
e-mail: geoff@bci.gb.com

Secretary

Clive Weeks
Tel: +44-(0)1252-674890
e-mail: Drcliveweeks@aol.com

Sponsorship Secretary

John Holliday
Tel: +44-(0)114-222-2685
e-mail: J.D.Holliday@sheffield.ac.uk

Newsletter Editor

Frances Daniel
Tel: +44-(0)1949-844837
e-mail: francesdaniel@bottesford65.fsnet.co.uk

Next Copy Date

January 31 2004

The Ernie Hyde Award through the Ages - Part III

Each year, the Chemical Structure Association gave an Award for an outstanding contribution to the Society, and the Trust intends to continue this tradition, in memory of Ernie Hyde who was a pioneer in the structure handling field in the 1960s and 1970s. It is customary, from time to time, to provide an updated list of all the Ernie Hyde award winners from the first Award in 1977 (for the original Chemical Notation Association). It seems appropriate once more to print the full list, following the news of the sad death of the first CSA recipient in 1983, Jane Whittall (née Gaworska), and also the news in 2002 from Ernie's daughter of the death of her mother, Barbara.

Research by Peter Nichols has revealed the following Award winners:

1977	Sandra Ward	1989	Ian Tarr
1978	Janet Clarke	1990	Phil McHale
1979	Barrie Walker	1991	Charles Citroen
1980	Edwin Kyte	1992	Arthur McArdle
1981	Pamela Chubb	1993	Steve Welford
1982	Phil McHale	1994	John Barnard
1983	Janet Ash (for CNA) and Jane Gaworska (for CSA)	1995	No award
		1996	No award
		1997	Andrew Poirrette
1984	Wendy Warr	1998	Peter Nichols
1985	Mike Elder	1999	Ian Bruno
1986	Mike Lynch	2000	Gez Cross
1987	Peter Baker (posthumously)	2001	Rob Brown and Rosemary Downs
1988	No award	2002	John Holliday

Third Joint Sheffield Conference on Chemoinformatics

The Chemical Structure Association Trust and the Molecular Graphics and Modelling Society announce their Third Joint Sheffield Conference on Chemoinformatics. The conference will be held in The Octagon Centre and Tapton Hall of Residence, University of Sheffield, UK, from 21st to 23rd April 2004.

Further details of the conference and registration information will be available on the conference website (http://www.cisrg.shef.ac.uk/shef2004) in the near future. For enquiries about exhibition space or event sponsorship, please contact Dr. Val Gillet, e-mail v.gillet@sheffield.ac.uk.